CHAPTER 6

6. CLASSICAL MODELS FOR INTERMOLECULAR INTERACTIONS

One of the important applications of MO calculations at the high level of theory is to build simpler yet accurate models for intermolecular interactions. In this chapter we will describe our efforts to construct classical models of intermolecular interactions based on *ab initio* calculations. According to the Hellman-Feynman theorem, forces on nuclei in the molecular system can be calculated classically from the charge density of the molecule. Therefore, when building a classical model one has to make sure that (a) wavefunction complies with the Hellman-Feynman theorem, and (b) electron density is reproduced by classical charge distribution to a good approximation. Classical charge distribution schemes in the form of different partial atomic charge separation methods are considered in the first Section. We found that the *ab initio* values on interaction energy in the urea chain dimer are best reproduced by Mulliken charges. These values are used in Section 6.2 to describe polarization effects in larger chain clusters. The modification of the wavefunction to satisfy the Hellman-Feynman theorem by optimizing centroid positions of each basis function is described in Section 6.3. The resulting charge distribution is significantly improved, so that the residual electric field on the nuclei in the optimized molecule vanishes.

Since the existing codes are not well suited to handle floating basis sets, these calculations present a computational challenge. Reducing the number of the basis functions N to the minimum ($N=N_\sigma/2+N_\pi$, where N_σ and N_π are the number of σ- and π-electrons, respectively) greatly reduces computational costs while maintaining built-in polarization flexibility of the basis set. This possibility is considered in Section 6.4. A semiempirical approach to optimizing parameters of this minimal floating basis set is also suggested. The ability of the wavefunctions in the form of minimal floating basis set to be exactly represented by N^2 point charges opens the possibility of building classical and combined models based on these wavefunctions.

Finally, alternatives to the electrostatic approach in description of the H-bonds are considered in Section 6.5. Linear correlations between H-bonding energy and monomer properties were found for the number of H-bonded systems. These are interpreted in Section 6.5 in terms of the donor-acceptor nature of H-bonds.

6.1 Applicability of various definitions for atomic point charges

Atomic charge is not a quantum-mechanical observable and therefore does not have a unique definition. Methods of defining atomic charge can be classified in four groups: (I) empirical fit, including electronegativity schemes; (II) population analysis; (III) fitting to electrostatic potential; (IV) systematic corrections of population atomic charges to fit experimental dipole moments.[1,2] Here we will consider only class II and III definitions.

The first and most widely used definition of atomic charge, based on the wave function, was suggested by Mulliken.[3] The arbitrary aspect of this definition (equal splitting of the overlap populations) was subject to criticism and gave rise to several improved schemes, including that of Löwdin[4] and natural population analysis (NPA) by Weinhold,[5] in which the basis set is orthogonalized and the overlap population vanishes. Another approach is to divide charge density in real, rather then in Hilbert space. Such schemes were suggested by Hiershfeld[6] (based on electron density for spherical atoms), and Bader[7] (based on topological analysis of the total electron density). There are also charge partition schemes especially designed to reproduce intermolecular interaction energies. This group is called potential derived charges,[8] and they are optimized to give the best fit to the distribution of *ab initio* electrostatic potential around the molecule. Another approach is to calculate the partial atomic charge from *ab initio* force acting on the nucleus in the external electric field, perpendicular to the molecular plane. These charges are called force derived charges and were reported to accurately reproduce intermolecular interaction energy.[9] Unfortunately, they are difficult to define for non-planar molecules. The projection of all multipole momenta from all overlap densities on the nearest expansion point (an atom, a bond centroid, etc.) defines distributed charges (and multipoles) as suggested by Stone.[10] Finally, the gradient of molecular dipole moment with respect to the coordinates of a given atom is a definition of the charge according to Cioslovski[11] (one of the properties of this definition is that calculated IR intensities are equal to *ab initio* predicted values).

We applied the charge definitions named above to describe a chain dimer of urea. Lowdin and Stone charges were calculated using GAMESS-UK; the other charges were obtained with GAUSSIAN 94 using options of the keyword Pop (Regular for Mulliken, NPA for Weinhold, ChelpG and MK for potential-derived charges). Force derived charges were calculated from forces exerted on nuclei in a finite electric field, orthogonal to the molecular plane. The results are shown in Table 6.1. Presumably, if the charge scheme is correct, replacement of one molecule in the cluster with a set of point charges does not change polarization of the other molecules.

Examination of the differences between the charges in the dimer and in the

Table 6.1. Comparison between different atomic charge partition schemes for the chain dimer of urea, all results are at HF/D95** level, charges are in a.u., interaction energy is in kcal/mol.

	Mulliken	Lowdin	NPA	Stone	ChelpG	MK	Force-derived
Atomic PC in the monomer							
O1	-0.50	-0.42	-0.82	-0.90	-0.78	-0.74	-0.54
C1	0.48	0.24	1.06	1.17	1.30	1.21	0.54
N1	-0.60	-0.32	-0.95	-0.57	-1.23	-1.19	-0.75
H1	0.31	0.20	0.42	0.23	0.54	0.51	0.35
H1'	0.29	0.20	0.41	0.21	0.45	0.45	0.36
difference between charges in dimer vs. monomer							
O1	-0.018	-0.017	-0.019	-0.025	-0.039	-0.040	-0.008
C1	-0.020	-0.001	0.001	0.011	0.089	0.077	-0.026
N1	-0.012	-0.004	-0.007	-0.008	-0.089	-0.051	0.017
H1	0.038	0.007	0.024	0.038	0.081	0.054	-0.022
H1'	-0.020	-0.010	-0.013	-0.016	-0.008	-0.010	-0.012
O2	-0.029	-0.019	-0.042	-0.045	-0.134	-0.091	0.031
C2	0.025	0.017	0.019	0.010	0.145	0.103	0.012
N2	0.003	0.010	0.008	-0.000	-0.055	-0.022	0.018
H2	0.010	0.005	0.007	0.009	0.019	0.014	-0.001
H2'	0.002	0.001	0.001	0.002	0.022	0.015	-0.005
% error of polarization by PC vs. polarization in dimer							
O1	72	292	107	37	86		
C1	34	604	110	48	47		
N1	151	-17	133	70	42		
H1	110	693	135	32	58		
H1'	68	214	109	26	180		
O2	139	111	20	462	61		
C2	71	33	1096	1262	48		
N2	121	46	-2730	-20426	29		
H2	74	52	999	-1120	64		
H2'	25	31	3912	3528	45		
mean	88	192	403	-2123	67		
Molecular HF stabilization in the presence of PC							
$\Delta E(12)$	-8.29	-31.57	-13.12	-14.25	-13.04		-7.88
$\Delta E(21)$	-7.86	-5.55	-10.75	-8.72	-12.94		-7.15
$\Delta E(av)$	-8.07	-18.56	-11.93	-11.48	-12.99		-7.51

monomers shows a very similar picture of molecular polarization. The only exception is the decrease of the force related charge on an O2 atom accepting an H-bond. This is counterintuitive and allows us to eliminate this charge scheme. The next block reports error of the PC model in representing polarization in the dimer. Here we compare the deformation of the charges on one monomer exerted by a PC set representing another monomer with deformation of *ab initio* charge in the dimer. We can see that the Stone and NPA schemes overestimate the charge deformations on the second monomer (H-bond acceptor) by an order of magnitude. By contrast, the Löwdin scheme overestimates the charge changes on the first monomer. The Pop=MK option is not available in GAUSSIAN in the presence of external point charges, but seems to give the results similar to Pop=ChelpG option in the absence of point charges.

Thus, we are left with Mulliken and with potential derived charge ChelpG. Both pass the last test, monomer stabilization energy in the presence of point charges representing another monomer. Both are reasonably symmetrical and close to HF interaction energy (slightly greater, probably due to sterical repulsion in the HF dimer). In the following Section we will use Mulliken charges.

6.2 Variation of atomic point charges upon molecular polarization simulates cooperative effects

As discussed in Chapter 2, polarizability can be introduced into the force field by assigning an induced dipole to the molecular center or to each atom or bond. We examined the possibility of obtaining the same result by a much simpler approach. In this approach each atomic charge q in the molecule varies quadratically upon the external electric field F at the position of this atom:

$$q(F) = q_0 + a \cdot F + b \cdot F^2$$

Parameters a and b are different for each of x, y, and z components of the field. The set of parameters and the charge q_0 in the absence of the external field are individual characteristics of atoms in molecules. We fit these parameters to HF/D95** Mulliken charges in external dipolar field of ±0.01 au. We used quadratic rather then linear dependence to describe symmetric molecules. For example, an external field perpendicular to the planar molecule will induce the same atomic charges, as the field in the opposite direction. Therefore, the linear component of the charge dependence *a* is zero for this configuration, and the charge deformation is described by quadratic component *b*.

We applied the model of variable atomic charges to the urea chain clusters considered in Chapter 5. Intramolecular geometry was fixed to that of a monomer and H-bonding distance to that of the dimer. Since the external field experienced by the atom in a cluster is created by charges from the other monomers, it is no longer uniform. As a result, the sum of modified atomic charges deviates from zero. In the spirit of Mulliken analysis, we divided this deviation equally among all atoms so that the molecules remain neutral.

The results are shown in Table 6.2. As one can see, ΔE_n for the hexamer is increased by 40%, compared to a 15% increase for the constant charge model and a 50% increase for the HF results. The dipole moments and total stabilization energies are underestimated (up to 25% for higher clusters), as often happens when Mulliken charges are used. We can conclude that polarization as described by this model

Table 6.2. Dipole moments μ (D), and interaction energies for chain clusters of urea using HF/D95**, constant point charges (PC) and variable point charges (VPC) methods.

N	μ, HF	μ, PC	μ, VPC	ΔE_n, HF	Δ, %	ΔE_n-ΔE_{def}	Δ, %	ΔE_n,PC	Δ, %	ΔE_n,VPC	Δ, %
1	4.8	4.0	4.0								
2	10.9	7.9	9.0	-7.96	100	-8.19	100	-5.07	100	-6.84	100
3	17.5	11.9	14.0	-10.12	127	-10.67	130	-5.58	110	-8.71	127
4	24.2	15.8	19.1	-11.00	138	-11.73	143	-5.73	113	-9.22	135
5	31.1	19.8	24.2	-11.13	140	-11.96	146	-5.79	114	-9.42	138
6	38.0	23.8	29.4	-11.41	143	-12.29	150	-5.82	115	-9.51	139

accounts for a large part of the cooperative effect in H-bonding. Further improvements of the variable charges model may be necessary before it can be incorporated in the empirical force fields.

6.3 Improving calculated molecular electric properties with floating gaussian basis sets

There is a theoretical possibility of reducing intermolecular interactions to Coulomb forces in accordance with the Hellman-Feynman theorem. However, this requires high quality monomer wavefunction, properly deformed by intermolecular interactions. The medium size atom centered basis sets widely used in MO calculations produce wavefunctions that are not compliant with the Hellman-Feynman theorem. To improve the wavefunction, a larger number of basis functions with high angular momenta is necessary. The alternative is optimization of the center coordinates for all basis functions rather than keeping them fixed to nuclear positions (so-called floating gaussian basis set). It was shown that wavefunctions built with floating gaussian (FG) basis sets satisfy the Hellman-Feynman theorem.[12] For this reason the Coulomb interaction energy of charges obtained in FG basis (even within Mulliken approximations) is much close to the total HF interaction energy, as was shown by Dannenberg, Simon and Duran.[13]

Any model that reduces continuous electron density distribution to point charges, suffers a penetration problem. The interaction energy between a charged particle and a charge distribution is determined by the charge in the inner area of this distribution, as the outer area has no effect. So the interaction becomes weaker and vanishes when the particle reaches the center of the charge distribution. This effect is neglected by the point charge model, unless explicit corrections are made (e.g., in the form of damping factors). To reduce the error, Dannenberg, Simon and Duran used point charges to represent one part of the system, and *ab initio* electron density to represent another part.

As compared to conventional double-ζ basis sets, FGs do not significantly improve the total energy. However, the bonds, angles, dipole moments and polarizabilities of small molecules (2-4 atoms) are shown to be much closer to HF limits in FG basis sets.[14]

When standard MO programs are used, the FG centroid positions are treated together with atomic coordinates and determined according to the variational principle. Unfortunately, the existing algorithms for molecular geometry optimization are not well suited for FG specifics, such as large energy change at small coordinate displacements. Optimization problems drastically increase with the size of the molecule. Despite many insistent attempts, we were not able to obtain the wavefunction for the urea ribbon dimer in floating D95** basis set. The results for the urea chain dimer are presented in Table 6.4.

In calculations of the urea chain dimer in standard D95** basis set positions for all atoms and basis functions were optimized separately (sets of 3 p-functions and 6 d-functions moved together respectively). One can see that optimization of the positions for basis functions decreases the dipole moment for monomer and dimer (bringing it closer to experimental value), as well as intermolecular interaction (by about 1 kcal/mol). To access the quality of the point-charge representation for FG basis function, we performed calculations on 3 levels: PC-PC interaction, PC-electric field created by charge density of the monomer, and PC-electric field created by the charge density of the monomer polarized in a dimer. To measure the values of the latter, we performed the SCF procedure for the dimer, then modified the geometry so that the second monomer was removed to infinity (actually, 900Å), and read in the density matrix from the checkpoint file without repeating the SCF procedure (keywords Density=Checkpoint, Guess=Only). Reading in the wavefunction instead of the density matrix with modified geometry (keywords Guess=(Read,Only)) did not give any meaningful results. The interaction energies in all the models are reported in Table 6.3. Since molecules 1 and 2 in the dimer are represented differently, we report the results in both ways, as well as the average values.

Comparing PC-PC interaction energy to the last column of Table 6.4 (dimerization energy of rigid monomers) we can conclude that conventional Mulliken charges significantly underestimate the interaction energy, even if molecular polarization is taken into consideration. FG-based monomer charges perform better, but only (polarized) dimer charges give interaction energy obtained at HF level. Surprisingly, the most accurate approximation (floating PCs for the dimer in the field of the polarized monomer) significantly overestimates the interaction. This is probably due to steric repulsions, which are not taken into account by the PC model. All the other models give results reasonably consistent with *ab initio* interaction energy.

Table 6.3. Interaction energy for the chain dimer of urea calculated in Mulliken PC model (conventional and fully floating D95** basis set parametrization).

	conventional			Floating		
	E(12)	E(21)	E(av)	E(12)	E(21)	E(av)
Monomer PC:						
PC-PC	-5.13	-5.13	-5.13	-5.42	-5.42	-5.42
PC-monomer field	-6.45	-7.63	-7.04	-7.51	-7.39	-7.45
PC-dimer field	-6.39	-8.46	-7.42	-8.34	-9.35	-8.85
dimer PC:						
PC-PC	-6.20	-6.20	-6.20	-8.34	-8.34	-8.34
PC-monomer field	-7.60	-7.82	-7.71	-9.84	-8.68	-9.26
PC-dimer field	-7.55	-8.69	-8.12	-10.93	-11.00	-10.96

The difference between 1-2 and 2-1 interactions can be considered a measure of systematic error for the model. The lowest difference (except for the PC-PC model, where it is 0 by design) is observed for our best approximation (FG-based dimer PCs in the dimer field), as well as for FG-based monomer PCs in the monomer field. This is a result of uniform representation for both parts of the dimer. The next lowest difference is observed in the conventional dimer PC in the monomer field model, and is most likely an interplay between underestimated polarizability and overestimated polarity, typical for the conventional HF method.

We can conclude that the use of Mulliken point charges obtained in the FG basis set for the dimer is advantageous compared to those of the conventional basis set. It is also worth noting, that the polarization mechanism described by the FG basis set is well suited for classical implementation. In fact describing polarizability using the charges harmonically oscillating around the centers of heavy atoms had been suggested.[15] In this model the displacement of the charge from its equilibrium position at the atomic center increases the total molecular energy but creates an atomic dipole, stabilizing the molecule in the external field. The value of the charge and the force constant describing its displacement are obtained by fitting to *ab initio* electrostatic potential distribution in the presence of electric field. Results were found encouraging, but needed adjustments in Van der Waals parameters. Also, the anisotropy of polarizability was not accounted for in this model.

In our opinion, fitting the oscillator parameters to the experimentally obtained partial atomic polarizabilities[16] would yield a more transferable model. However, using the positions and populations of FG functions directly to describe the oscillating charge would eliminate the necessity of any fitting.

6.4 Construction of minimal floating spherical gaussian basis set wavefunction, and its exact point charge model

A simplified version of the FG basis set, in which only s-functions are used (floating spherical gaussian orbitals, FSGO) and both their positions and exponent parameters are optimized, has been suggested[16] and implemented.[17] It was shown that electric properties as well as energies are drastically improved in this approach for the whole range of interatomic distances from equilibrium to the dissociation limit.[18] More accurate multiconfigurational correlation treatment is also possible. As the orbitals are explicitly localized in FSGO basis set, we should expect to enjoy all the benefits reported for use of natural localized orbitals for this purpose.[19]

Spherical functions allow for analytical description of nonbonding intermolecular interactions without resorting to empirical parameter fitting. It was shown that orbital-orbital dispersion coefficients and three- and four-body non-additive corrections to the dispersion interaction energy are easily obtainable for these FSGO wavefunction using second and higher order perturbation energy.[20] Damping functions and exchange repulsion energy formulas were obtained using the surface integral method.[21]

Since the product of two gaussian s-functions

$\chi_i(\mathbf{r}) = (\alpha_i/\pi)^{3/2} \exp(- \alpha_i (\mathbf{r}-\mathbf{r}_i)^2)$; $\chi_j(\mathbf{r}) = (\alpha_j/\pi)^{3/2} \exp(- \alpha_j (\mathbf{r}-\mathbf{r}_j)^2)$

is an s-function itself,

$\chi_{ij}(\mathbf{r}) = \chi_i(\mathbf{r})\, \chi_j(\mathbf{r}) = (\alpha_{ij}/\pi)^{3/2}\exp(-\alpha_{ij} (\mathbf{r}-\mathbf{r}_{ij})^2)$; $\alpha_{ij} = (\alpha_i+\alpha_j)/2$; $\mathbf{r}_{ij} = (\alpha_i\, \mathbf{r}_i+\alpha_j\, \mathbf{r}_j)/(\alpha_i+\alpha_j)$.

The electron density of the FSGO wavefunction is a sum of N spherical gaussians, and can be exactly represented by the sum of (N^2-N) point charges.[22] To account for penetration in close proximity to the molecule, dumping factors must be used. The

FSGO model was modified to describe *p*-electrons using a tetrahedral arrangement of FSGO around the core[23] or lobe functions, i.e., pairs of s-orbitals at a fixed distance to simulate *p*-orbital.[24] The latter were found successful in reproducing geometries and energies of hydrides HX and Van der Waals dimers, including Ar...HCl.[25] The former yielded fairly small errors in bond length (within 2%) and conformational barriers (but not the total energies) for series of hydrocarbons, and allows one to establish trends in positions and exponents of FSGO, depending on local environment.[26] Simple geometric rules were found to predict the size and approximate location of the FSGO centers based on closest packing principle:[27] each orbital is assigned a radius based on its exponent factor, so that these spheres are touching each other in the molecule but do not penetrate. Once the rules for the molecules of a certain class are established, these parameters may be treated as constants to allow HF as well as post-HF calculations on large molecules and clusters.[28]

As one can see, HF/FSGO is a powerful yet computationally light *ab initio* method, capable of quantitative treatment of intermolecular interactions. The numerical optimization of exponent parameters is clearly a disadvantage, and has

Table 6.4. Comparison between chain dimer of urea optimized with at HF level with conventional and fully floating basis set: dipole moments μ (D), total energy E (a.u.), interaction energy, CP correction, CP-corrected interaction energy, and CP-corrected interaction energy less monomer relaxation (kcal/mol).

method	monomer μ	dimer μ	E	ΔE	ΔEcp	ΔE'cp
HF/6-311+G(3df,2p)	4.59	10.61	-448.173633	-7.67	-7.50	-8.08
conventional HF/D95**	4.70	10.84	-448.106299	-8.78	-8.08	-8.44
floating HF/D95**	4.61	10.70	-448.121097	-8.02	-7.01	-8.16
HF/FSGO, model 1	5.59	11.65	-381.296431	-9.62	-6.66	-7.89
HF/FSGO, model 2	6.05	13.06	-393.211024	-11.81	-9.45	-10.84
HF/FSGO, model 3	2.98	6.84	-385.459652	-4.62	-2.05	
HF/FSGO, model 4	4.80	10.69	-381.776555	-8.98	-6.04	

prevented this method from being widely used.

We are mostly interested in the FSGO model as an inexpensive way to obtain a wavefunction with accurate electric properties. Hence, it would be logical to select exponent parameters of FSGO functions, which reproduce electrostatic potentials obtained at HF or correlated limits, rather than from the variational principle. This could be achieved by a semiempirical adjustment of the optimal exponent factors for small molecules or functional groups. The values obtained could be then tabulated similarly to conventional basis sets and used for complete HF/FSGO optimizations of large clusters and macromolecules. Since FSGO wavefunctions can be exactly represented by classical force fields (PC model with steric and dispersion terms), total optimization can be carried out on a fragment-by-fragment basis, allowing a high degree of parallelization. This has the advantage of a built-in mixed QM/MM technique, where *ab initio* fragments can be as small as an isolated bond or a lone pair.

As a first step toward this goal, we carried out preliminary FSGO calculation on the urea chain dimer. First, the parameters of the basis set were obtained in atomic calculations on low-spin states of C and O atoms. In these calculations one atom-centered spherical function described the core and the octahedron of six spherical functions (three lobe functions) around it described the valence shell. The distance from the lobe functions to the atomic center (separation distance) was varied to fit the quadrupole moments, while all exponents were optimized using GAMESS-US. The calculations yielded separation distances of 0.20 and 0.55Å, with exponent factors of 0.561 and 0.571 for O and C, respectively. Other atoms were nearly spherical and gave a good fit at a large interval of the parameter. For F and N we chose separation distances of 0.12 and 0.30Å, which corresponded to the optimal exponent factor of 0.56 (to keep basis set uniform). The exponent factors for the central core function were almost independent of separation distance and were found to be 10.58, 13.45, and 18.23 for C, N, and O atoms.

In molecular and dimer calculations, core functions were placed on C, N, and O atoms, one or two lobe functions were retained on atoms O or N respectively, and single spherical functions were used to describe σ-lone pair on O atom, NH, CN, and CO bonds. The exponent values were fixed at value 0.56, obtained in atomic calculations for all valent SGOs (model 1). All geometrical parameters (atomic and electronic, except lobe separations) were optimized using GAUSSIAN 98. Unfortunately, optimization with the exact Hessian does not work for floating basis sets. That is why the options for the keyword OPT were set to Tight, Z-Matrix, NRScale, EstmFC and Iop(3/15=0) for scaled gradient optimization using small steps. The bond lengths and angles, listed on Table 6.5, were fairly close to conventional HF results. Interaction energies and dipole moments for the dimers and monomers are listed in Table 6.4. One can see that H-bond energy after CP correction closely matches high-level HF results, even though the dipoles are clearly overestimated.

BSSE in model 1 was found to be rather large (about 3 kcal/mol). In an attempt to reduce this error, we used a more realistic description of the core orbitals in the form of six contracted gaussians, taken from D95 basis set (model 2). This decreased the total energy by 12 Hartree, but CP correction decreased only by 20%. Due to high gradients, we were not able to optimize the positions of the core functions, and fixed

Table 6.5. Bond lengths and angles for urea monomer and H-bond in the chain dimer

	C=O	C-N	N-Hs	N-Ha	OCN	CNHs	CNHa	H...O
HF/6-311+G(3df,2p)	1.196	1.358	0.988	0.989	122.3	117.4	123.4	2.26
FSGO, model 1	1.292	1.396	1.001	1.001	126.8	120.4	123.2	1.98
FSGO, model 2	1.238	1.456	1.050	1.045	128.2	119.1	121.9	2.19
FSGO, model 3	1.331	1.375	1.001	1.001	126.8	120.4	123.2	2.01
FSGO, model 4	1.218	1.354	0.957	0.956	127.3	121.1	121.6	1.92

them on the nuclei. This further increased dipole moments and H-bonding interaction, and did not improve molecular geometry.

Next, the exponent factors for all SGO in the monomer were individually optimized using G98OPT utility. The resulting model (model 3) stabilizes the total energy by 2 Hartree, but strongly underestimates dipole moments (Table 6.4) and, as a result, intermolecular interaction. This supports our idea to choose exponents which well reproduce electrostatic properties, rather than minimize the total energies. To illustrate this, we fixed exponents of lobe functions at atomic values and varied exponents of in-plane lone pairs on O atoms, as they were found to have the greatest impact on the dipole moment. The dipole moment, which is reasonably close to the high-level HF calculation, was obtained with the exponent factor value of 0.60 (model 4). This significantly improved bond lengths (they are within 0.03Å of high-level HF values). Uncorrected interaction energies decreased by 0.7 kcal, and CP correction did not change.

We can conclude that the FSGO model is precise enough to describe intermolecular interactions and, upon standardization of the technique and algorithm improvements, may well compete in accuracy with conventional HF calculations. On the other hand, low computational cost and the possibility of exact classical representation of FSGO wave function argue for its use in the development of presumably more precise hybrid quantum mechanics/molecular mechanics (QM/MM) methods and "on the fly" molecular dynamics simulations.

6.5. Non-electrostatic models for intermolecular interactions

The earlier attempts to interpret mutual orientation of the molecules in crystals and gas phase dimers were based on the classical Lewis structure (formulated by Bent[29] as the electron pair close packing principle) and were reasonably successful. Later developments went in two directions: orbital models, based on localization procedures of various flavors, and reducible down to maximum overlap considerations;[30] electrostatic approach,[31] based on general domination of the electrostatic term in the total interaction energy.[32] As follows from the Hellman-Feynman theorem, orbital and electrostatic approaches should be equivalent. In fact, a model in which the point charges are placed at the centroids of localized orbitals was suggested by Kollman[33] to predict the geometry of molecular complexes.

Choosing the "best" approach should be a matter of convenience and simplicity. One way is to choose the approach offering more predictability at lower cost, rather than the one, which is better justified theoretically. In this respect the FSGO model, localized by design, is intuitively clear, and yet can easily be quantified from both the orbital and the electrostatic viewpoints.

There are a few examples in the literature of non-electrostatic qualitative description of intermolecular interactions in the literature. One is a linear correlation between the acidity/basicity of the hydrogen bond donor and acceptor.[34] In the case of small O-containing compounds, H-bonding with water varied despite the similarity of the charge on the O atom, whereas for substituted amines the charge significantly differred, and H-bonding strength remained almost the same.[35] The authors noticed that H-bonding energy correlates more strongly with the acidity or basicity of the participating groups than with their partial charges. An opposite conclusion was made in the case of H-bonded complexes between substituted acetylides and methanol, R-$C\equiv C^-$...$HOCH_3$ (R=H, *tret*-Bu, Ph, *para*-$PhCH_3$), based on experiments in the gas phase.[36] These authors found the same H-bond strength (21.5 kcal/mol) for all four complexes, despite different basicity of acetylides (in 8 kcal/mol range). Based on potential derived charges obtained at MP2/6-311++G** level, they argued that electron distribution in $C{\equiv}C^-$ fragment (as well as in C≡C-H fragment[37]) is independent on the substituent.

We looked at this relationship quantitatively on the example of *para*-substituted phenylacetylenes (Table 6.6). Enthalpies of deprotonation (DPE) and H-bond formation (HBE) were calculated with AM1 Hamiltonian. We found an excellent correlation (R^2=0.997) for the enthalpy (Figure 6.1):

$$\text{HBE} = -16.27(\pm 0.05)\ \text{kcal/mol} + 0.0375(\pm 0.0007) \times \text{DPE}$$

and less pronounced correlation (R^2=0.970) for the H-bonding distance:

$$r(\text{H...O}) = 1.865(\pm 0.004)\ \text{Å} + 8.9(\pm 0.5)\ \text{Å mol/kcal} \times 10^{-4}\ \text{DPE}$$

Table 6.6. AM1 results for *para*-substituted phenylacetylenes X-Ph-C≡C-H (A): enthalpy of formation ΔH_f for the monomers, their anions, and complexes with water; enthalpies of deprotonation and H-bonding (kcal/mol), bond lengths (Å), and H-boning angles (^{o}).

X=	ΔH_f(AH)	$\Delta H_f(A^-)$	ΔH_f(AH...OH_2)	ΔΔH(deprot)	ΔΔH(H-bond)	C-H	H...O	CH...O
NMe_2	85.05	104.90	24.09	385.53	-1.72	1.065	2.2079	166.1
NH_2	74.50	94.04	13.51	385.22	-1.75	1.065	2.2149	168.0
H	76.49	94.34	15.36	383.53	-1.89	1.065	2.2017	166.3
OH	32.07	49.50	-29.06	383.15	-1.88	1.065	2.2047	169.6
F	31.27	45.27	-30.01	379.68	-2.04	1.065	2.2082	180.
Cl	69.43	82.68	8.12	378.93	-2.07	1.065	2.2072	180.
CN	108.19	116.23	46.67	373.72	-2.28	1.065	2.2006	180.
NO_2	80.35	81.24	18.53	366.57	-2.58	1.065	2.1894	180.
NH_3^+	231.22	173.59	167.30	308.05	-4.68	1.065	2.1407	180.

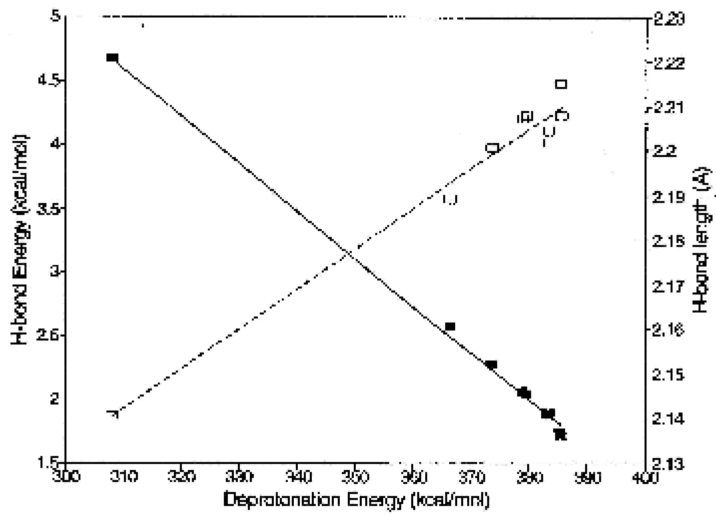

Figure 6.1. Linear correlations for complexes of substituted *para*-phenylacetylenes with water.

It is interesting to note that in the weakest complexes (X=NH_2, OH, H) the fragment H_2O...H are not planar. This feature is not easy to rationalize based on the electrostatic model. Here we suggest an explanation based on MO description of H-bond. Let us consider the interaction between electron acceptor orbital *a* localized on C-H fragment with electron donor orbitals n_σ, n_π, representing lone pairs of an O atom. According to second-order perturbation theory, the interaction between two orbitals is inversely proportional to the difference in their energy levels ϵ and proportional to the square of the Hamiltonian integral. Habitual replacement of the Hamiltonian with the overlap integral yields interaction energy

$$\Delta E = - k <a|n_\sigma>^2/(\epsilon(a)-\epsilon(n_\sigma)) - k <a|n_\pi>^2/(\epsilon(a)-\epsilon(n_\pi))$$

where k is proportionality constant, the first and second terms represents a-n_σ, and a-n_π interaction. Keeping the H...O distance constant, we will look at the dependence of ΔE on the angle α between the Z axis (C_2 axis of water molecule) and H...O direction. Assuming the lone pairs of the O atom to be sp_z^n and p_x-AOs, we can express the angular dependence of the overlap integral as:

$<a|n_\sigma> = 1/\sqrt{(n+1)}<a|s> + \sqrt{n}/\sqrt{(n+1)}<a|p_z>=$

$\sqrt{(n+1)}<a|s>+\sqrt{n}/\sqrt{(n+1)}<a|p>\cos\alpha$

$<a|n_\pi> = <a|p_x> = <a|p> \sin\alpha$

In the case of pure s-character of the lone pair, n=0 and the angular dependence becomes constant. For the angular dependence of interaction energy, we now have:

$$\Delta E(\alpha) = - k\,(1/(n+1)\ <a|s>^2 + n/(n+1)\ <a|p>^2\cos^2\alpha + 2\sqrt{n}/(n+1)\ <a|p>\ <a|s>\cos\alpha)/(\epsilon(a)-\epsilon(n_\sigma)) - k\ <a|p>^2\ (1-\cos^2\alpha)/(\epsilon(a)-\epsilon(n_\pi))$$

Taking the derivative of the interaction energy with respect to $\cos\alpha$ allows us to search for extremum:

$$\delta\ \Delta E/\delta(\cos\alpha) = - k\,(2n/(n+1)<a|p>^2\cos\alpha + 2\sqrt{n}/(n+1)\ <a|p>\ <a|s>)/(\epsilon(a)-\epsilon(n_\sigma)) + 2k\ <a|p>^2\cos\alpha/(\epsilon(a)-\epsilon(n_\pi)) = 0$$

which yields:

$$(n/(n+1)\ <a|p>\cos\alpha + \sqrt{n}/(n+1)\ <a|s>)/(\epsilon(a)-\epsilon(n_\sigma)) = <a|p>\cos\alpha/(\epsilon(a)-\epsilon(n_\pi))$$

or

$$\cos\alpha = \sqrt{n}/(n+1)\ <a|s>\ /\ <a|p>\ /\ (\ (\epsilon(a)-\epsilon(n_\sigma))/(\epsilon(a)-\epsilon(n_\pi)) - n/(n+1)\ )$$

We can substitute values obtained at HF/D95** level. The ratio of overlap integrals is <1sH|2sO>/<1sH|2pO> = 0.08/0.15 = 0.56 at internuclear distance of 2Å, and 0.011/0.027=0.40 at 3Å. Lone pair energies for H_2O $\epsilon(n_\sigma)$=-0.57 au, $\epsilon(n_\pi)$=-0.50 au, and antibonding energies for HF and HCN molecules are ϵ(H-C)=0.41 and ϵ(H-F)=0.22 au. Assuming n=1, $\cos\alpha$ = ½ 0.40/ (1.08-1/2) =0.12 for HCN, and ½ 0.56/ (1.10-1/2) = 0.34 for HF. This yields $\alpha=29^o$ and 33^o for HCN and HF respectively (compare to HF/D95** optimized values of 0^o and 50^o). So the angular dependence on overlap and energy is not very strong. The angle, however, strongly depends on the change in the hybridization state of the lone pair. As hybridization of σ-lone pair changes from sp^1 to $sp^{1.41}$, the angle between the C_{2v} axis of the H_2O molecule and the direction of the O...H bond decreases from 29^o to 1^o. Upon a further increase in p-character smooth minimum in the angular dependence becomes a singularity extremum at 0^o. The hybridization state of the lone pair can be estimated, for instance, in NBO analysis. We found it to be 1.05 for weak H-bonded complexes and free water molecules and 0.88 for the complex with HF. This brings the optimal angle values to 27^o and 37^o, accordingly. The hybridization state is, of course, model-dependent, but the trends should be similar regardless of the model used.

Another correlation for H-bond energy was found in a recent HF/6-31+G** study of XCHO...HF and XCN...HF complexes.[38] A common linear relationship was found between H-bonding energy and electrostatic potential at the position of H-acceptor atom O or N in the isolated molecule (rather than in the region of the H-bond). In the same study, no satisfactory correlation was found between H-bonding energy and Mulliken or potential derived charges on an O atom.

Electrostatic potentials on nuclei are not often considered. Among the few examples of their use, we could name the analysis of intramolecular interactions[39] and as a basis to derive a set of atomic charges.[40] The experimental information about electrostatic potential on the nuclei in principle can be extracted from X-Ray photoelectron spectroscopy.[41] Additional information on electric field gradient on the nuclei is available using microwave spectra,[42] NMR relaxation data,[43] and nuclear quadrupole resonance.[44] Effects of H-bonding on quadrupole tensor were studied recently on the examples of urea and HCN.[45]

We compared potential-HBE correlation with DPE-HBE correlation on the example of X-C≡C-H...N≡C-Y complexes (Figure 6.2). Mulliken charge and potential on atoms forming H-bonds, enthalpy of deprotonation, and H-bonding were again calculated at AM1 level (Table 6.7). All geometric parameters were optimized for

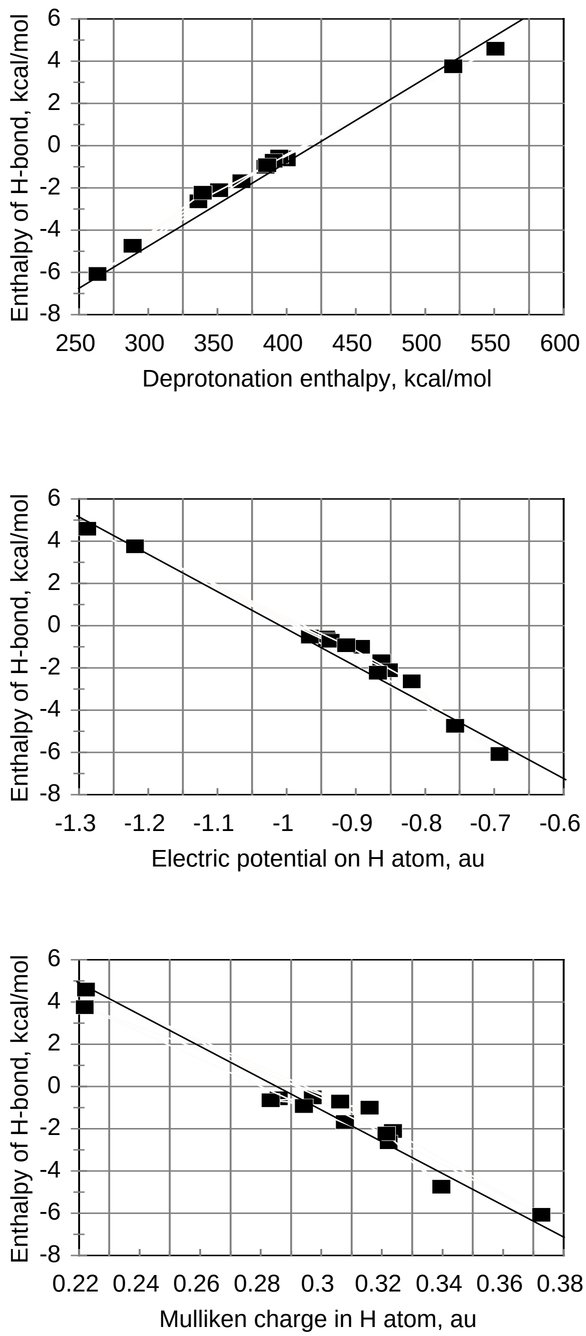

Figure 6.2. Linear correlations for complexes of substituted acetylenes with HCN.

Table 6.7. Complexes of substituted acetylene with substituted cyanide (AM1 results): H-bonding enthalpy HBE, protonation and deprotonation enthalpies PE, DPE, kcal/mol, electrostatic potential at the nucleus position in the monomer, and Mulliken charge on H-bonding atoms in the monomer.

	R-C≡N...H-C≡C-NO2				H-C≡N...H-C≡C-R			
R =	HBE	PE	ϕ(N)	q(N)	HBE	DPE	ϕ(H)	q(H)
O^-	-16.50	-336.51	-5.361	-0.534	4.59	551.04	-1.288	0.222
S^-	-12.86	-332.11	-5.233	-0.390	3.75	520.39	-1.219	0.222
NH_2	-2.39	-189.80	-4.882	-0.100	-0.52	394.80	-0.966	0.297
CH_3	-2.36	-189.07	-4.841	-0.095	-0.57	397.40	-0.943	0.286
H	-2.12	-182.08	-4.832	-0.105	-0.66	400.20	-0.942	0.283
OH	-1.97	-177.04	-4.837	-0.079	-0.72	390.93	-0.936	0.306
SH	-1.86	-182.90	-4.798	-0.037	-0.92	386.05	-0.914	0.294
F	-1.51	-165.78	-4.764	-0.047	-1.00	384.96	-0.892	0.316
CF_3	-0.90	-165.17	-4.713	0.036	-1.68	367.54	-0.863	0.308
NO_2	-0.44	-160.28	-4.705	0.068	-2.12	351.41	-0.852	0.324
$C(NO_2)_3$	-0.35	-159.26	-4.708	0.098	-2.24	339.60	-0.868	0.322
SO_2F	-0.04	-158.77	-4.660	0.137	-2.65	336.48	-0.819	0.322
$CH_2NH_3^+$	2.97	-88.51	-4.594	0.111	-4.75	289.01	-0.756	0.340
NH_3^+	4.22	-54.47	-4.490	0.228	-6.09	263.62	-0.692	0.373

most dimers. However, anionic acetylenes (X=O^-, S^-) and cationic cyanides (Y=NH_3^+, $CH_2NH_3^+$) were found unbound, and the N...H distance in these cases was fixed at 2.8 Å, a value obtained for the weakest H-bonds.

For Y=H and different substituents X in the donor molecule (Figure 6.3), the best correlation (R^2=0.991) was again with the basicity:

$$HBE = -15.0(\pm 0.3)\ \text{kcal/mol} + 0.036(\pm 0.001) \times DPE$$

Electric potential (ESP_H) on an H atom is also correlated (R^2=0.970) with H-bonding enthalpy:

$$HBE = 17.0(\pm 0.5)\ \text{kcal/mol} - 17.2(\pm 0.9)\ \text{kcal/mol/au} \times ESP_H$$

and Mulliken charge (Q_H) on H atom is correlated poorly (R^2=0.954):

$$HBE = 9.0(\pm 0.6)\ \text{kcal/mol} - 67.(\pm 4.)\ \text{kcal/mol/au} \times Q_H$$

For X=NO_2 and different substituents at H-bond acceptor we found rather poor correlation (R^2=0.953) with protonation energy PE of the cyanide:

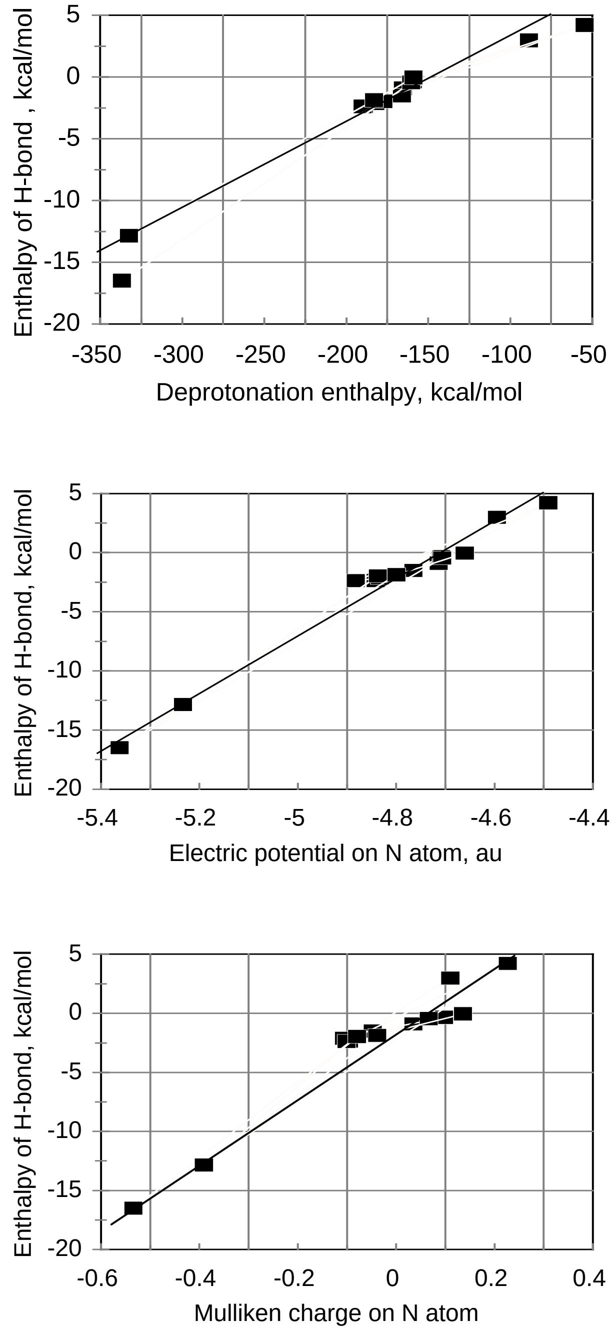

Figure 6.3. Linear correlations for complexes of substituted cyanides with nitroacetylene.

$$HBE = 10.(\pm 1.)\ kcal/mol + 0.07(\pm 0.005) \times PE,$$

whereas the correlation with electric potential on N atom was the best (R^2=0.974):

$$HBE = 111.(\pm 1.)\ kcal/mol + 23.(\pm 1.)\ Kcal/mol/au \times ESP_N$$

and the Mulliken charge on the N atom was poorly correlated again (R^2=0.953):

$$HBE = -1.(\pm 1.)\ kcal/mol + 26.(\pm 2.)\ kcal/mol/au \times Q_N$$

As ESP was the only quality to correlate with H-bonding strength for both donor and acceptor atoms, we can combine this property of donor and acceptor in a double correlation:

$$HBE = 94.5\ kcal/mol + 23.\ kcal/mol/au \times ESP_N - 17.\ kcal/mol/au \times ESP_H$$

The relationship can be used to predict H-bonding enthalpies for the complexes not included in the training set. For example, H-bond in H-C≡C-H...N≡C-H complex is predicted to be -0.62 kcal/mol, while AM1 calculation gives -0.66 kcal/mol.

A possible explanation for the linear dependence between H-bonding energy and electric potential on the nucleus could be found in a localized orbital picture of H-bonding. The energy of the lone pair on the H-acceptor atom, as well as the energy of the C-H antibonding orbital (which is mostly 1s orbital on H atom) are directly proportional to the electric potential on these atomic centers. In second order perturbation treatment the interaction energy between two orbitals is inversely proportional to the difference in their energy (assuming orbital overlap does not change). Fourier series expansion of the inverse proportionality gives the linear dependence in the second term. It would be interesting to build a potential function of the H-bond based on this relationship.

We can conclude that, at least for the complexes considered, electrostatic potential on atomic position in an isolated molecule can be used to evaluate its ability to serve as an H-bonding donor or acceptor.

References of Chapter 6

[1] Storer, J.W.; Geiesen, D.J.; Cramer, C.J.; Truhlar, D.J.: J. Comput.-Aided Mol. Design, **1995**, 9, 85.

[2] Li, J.; Zhu, T.; Cramer, C.J.; Truhlar, D.G.: J. Phys. Chem.A., **1998**, 102(10), 1820.

[3] Mulliken, R. S.: J. Chem. Phys. **1955**, 23, 1833.

[4] Lowdin, P.-O.: Phys. Rev. **1954**, 97(6), 1474.

[5] Reed, A.E.; Weinstock, R.B.; Weinhold, F.: J. Chem. Phys. **1985**, 83(2), 735.

[6] Hirshfeld F.L.: Theor. Chim. Acta, **1977**, 44, 129.

[7] Bader, R.F.W. Atoms in Molecules: A Quantum Theory, **1990,** Oxford Univ. Press: Oxford.

[8] Besler, B.H.; Merz, K.M.Jr.; Kollman, P.A.: J. Comp. Chem. **1990,** 11, 431.

[9] Dinur, U.: J. Comp. Chem. **1991**, 12, 469.

[10] Stone, A.J.; Chem. Phys. Lett., **1981**, 83, 233.

[11] Cioslovski, J.: J. Am. Chem. Soc., **1989**, 111, 8333.

[12] Hurley, A.C., Proc. R. Soc. London, Ser. A, **1954,** 226, 170, 179, 193.

[13] Dannenberg, J. J., Simon, S., Duran, M., J. Phys. Chem A; **1997**; 101(8); 1549.

[14] Helgaker, T., Almlöf, J., J. Chem. Phys**., 1988,** 89, 4889.

[15] Lamoureux, G.; Allouche, D.; Souaille, M.; Roux B.: Biophisical Journal, **2000**, 78 (1, part 2), 330A.

[16] Boys, S.F.; Proc. Roy. Soc., **1950**, A200, 542.

[17] (a) Frost, A.A. J. Chem. Phys., **1967**, 47, 3707; (b) Frost, A.A.; Rouse, R.A. J. Am. Chem. Soc., **1968**, 90, 1968.

[18] Tachikawa, M.; Taneda, K.; Mori, K.: Int. J. Quant. Chem. **1999**, 75, 497.

[19] Nemukhin, A. V.; Weinhold, F.: J. Chem. Phys. **1992**, 97(2), 1095.

[20] (a) Amos, A.T.; Yoffe, J.A. Theoret. Chim. Acta, **1976**, 42, 247; (b) Amos, A.T.;

Yoffe, J.A. Chem. Phys. Lett., **1976**, 39, 53.

[21] Kleinekathofer, U.; Tang, K.T.; Toennies, J.P.; Yiu, C.L.: J. Chem. Phys., **1997**, 107(22), 9502.

[22] Hall, G.G. Chem. Phys. Lett., **1973**, 20, 501.

[23] Chu, S.Y.; Frost, A.A. J. Chem. Phys., **1971**, 54, 760.

[24] Shih, S.; Buenker, R.O.; Peyerimhoff, S.D.; Wirsam, B.: Theoret. Chim. Acta, **1970**, 18, 277.

[25] Blustin, P.H.: Theoret. Chim. Acta, **1978**, 47, 249.

[26] Blustin, P.H.; Linnett, J.W.: J. Chem. Soc. Faraday II, **1974**, 70, 247.

[27] Blustin, P.H.: Chem. Phys. Let., **1975**, 35(1), 1.

[28] Maggiora, G.M; Petke, J.D.; Christoffersen, R.E.: Theor. Models Chem. Bonding, **1991,** 4, 65.

[29] Bent, H.A,: J. Chem. Educ., **1963**, 40, 446.

[30] Weinhold, F.: J. Chem. Educ., **1999**, 76(8), 1141.

[31] Buckingham, A. D.; Fowler, P. W. J. Chem. Phys. **1983**, 79, 6426.

[32] Umeyama, H.; Morokuma, K.: J.Am.Chem.Soc., **1977**, 99, 1361.

[33] Kollman, P.: J. Am. Chem. Soc., **1977**, 99(15), 4875.

[34] Hibbert, F.; Emsley, J.: In Advances in Physical Organic Chemistry; Bethell, D., Ed.; Academic Press: London, **1990**; Vol. 26.

[35] Marten, B.; Kim,K.; Cortis, C.; Friesner, R.A.; Murphy, R.B; Ringnalda, M.N.; Doree Sitkoff, D.; Honig, B.**:** J. Phys. Chem., 100 (28), 11775.

[36] Chabinyc, M.L.; Brauman J.I.: J. Phys. Chem. A; **1999**; 103(46); 9163.

[37] Wiberg, K.B.; Rablen, P.R.: J. Am. Chem. Soc., **1993**, 115, 9234.

[38] Galabov, B.; Bobadova-Parvanova, P.: J. Chem. Phys., **1999**, 103(34), 6793.

[39] Fliszar, S.; Vauthier, E.C.; Barone, V.: Adv. Quantum Chem. (**1999**), 36, 27.

[40] Su, Z.: J. Comput. Chem. (**1993**), 14(9), 1036.

[41] Chong, D.P.; Hu, C.-H.; Duffy, P.: Chem. Phys. Lett., **1996**, 249(5,6), 491; De Brito, A.N.; Correia, N.; Svensson, S.; Ågren, H.: J. Chem. Phys., **1991**, 95(4), 2965.

[42] Aldrich, P.D.; Kukolich, S.G.; Campbell, E.J.: J. Chem. Phys., **1983**, 78(6), 3521; Leung, H.O.: J. Chem. Phys., **1997**, 107(7), 2232; Howard, N.W.; Legon, A.C.; J. Chem. Phys., **1988**, 88(8), 4694.

[43] Cummins, P.L.; Bacskay, G.B.; Hush, N.S.; Halle, B.; Engström, S.: J. Chem. Phys., **1985**, 82(4), 2002.

[44] Ivanov, A.I.; Rebane, T. K.: Opt. Spectr., **1991**, 71(2), 146; Sham, T. K.: J. Chem. Phys., **1979**, 71(9), 3744.

[45] Aray, Y.; Gatti, C.; Murgich, J.: J. Chem. Phys., **1994**, 101(11), 9800; King, B.F.; Farrar, T.C.; Weinhold, F.: J. Chem. Phys., **1995**, 103(1), 348.